SSR Journal of Artificial Intelligence (SSRJAI)

OPEN ACCESS

Homepage: https://ssrpublisher.com/ssrjai/

Volume 2, Issue 5, 2025 ISSN: 3049-0413

Neural Networks and Deep Learning Models in Intrusion Detection Systems

I.O. Clementina[1], Dr. E.S. Chaku[2], Dr. S. Bassey[3], Dr. Emomotimi Agama[4], Prof. Gilbert Aimufua[5] & M.A. Ya'a[6]

[1]Master's Degree Student, Centre for Cyberspace, Department of Cybersecurity, Nasarawa State University, Keffi, Nigeria

[2]PG Coordinator, Centre for Cyberspace, Department of Cybersecurity, Nasarawa State University, Keffi, Nigeria

[3]Lecturer, Centre for Cyberspace, Department of Cybersecurity, Nasarawa State University, Keffi, Nigeria

[4]Director General, Security and Exchange Commission, Federal Republic of Nigeria

[5]Director, Centre for Cyberspace, Department of Cybersecurity, Nasarawa State University, Keffi, Nigeria

[6]PhD Student, Centre for Cyberspace, Department of Cybersecurity, Nasarawa State University, Keffi, Nigeria

Received: *25.08.2025* | ***Accepted:*** *21.09.2025* | ***Published:*** *25.09.2025*

***Corresponding author:** I.O. Clementina

DOI: *10.5281/zenodo.17357022*

Abstract **Original Research Article**

Intrusion Detection Systems (IDS) are critical for safeguarding modern networks against increasingly sophisticated cyber threats. Traditional IDS approaches, often signature-based, struggle to detect novel or evolving attacks, highlighting the need for intelligent, adaptive mechanisms. Neural networks and deep learning models have emerged as promising solutions due to their ability to learn complex patterns and generalize from large datasets. This study explores the integration of neural network architectures—including Multi-Layer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks—into IDS frameworks to enhance detection accuracy and reduce false positives. We provide a comprehensive analysis of supervised, unsupervised, and hybrid learning approaches, examining their performance in identifying diverse attack types such as Denial-of-Service (DoS), probe attacks, and insider threats. The study also addresses challenges in applying deep learning to IDS, including data imbalance, high dimensionality, feature selection, and real-time processing constraints. Comparative evaluations demonstrate that deep learning-based IDS consistently outperform traditional machine learning methods, particularly in detecting zero-day attacks and complex multi-stage intrusions. Additionally, we discuss the role of autoencoders, Generative Adversarial Networks (GANs), and reinforcement learning in enhancing IDS adaptability and resilience. The findings underscore the potential of deep learning to transform IDS into proactive, intelligent security solutions capable of continuous learning and adaptation in dynamic network environments. Future research directions include optimizing model interpretability, reducing computational overhead, and developing standardized benchmarks to evaluate deep learning-based IDS performance across heterogeneous network scenarios.

Keywords: Intrusion Detection System, Neural Networks, Deep Learning, Multi-Layer Perceptron, Convolutional Neural Network, Recurrent Neural Network, Long Short-Term Memory, Cybersecurity, Zero-Day Attacks, Feature Selection.

1. INTRODUCTION

Intrusion Detection Systems (IDS) have been pivotal in safeguarding computer networks from unauthorized access and malicious activities. Traditionally, IDS have been categorized into two primary types: signature-based and anomaly-based systems. Signature-based IDS detect known threats by matching patterns against predefined signatures, while anomaly-based IDS identify deviations from established network baselines, potentially flagging novel or unknown attacks.

However, with the increasing sophistication of cyber threats, traditional IDS methods have shown limitations. They often struggle to detect zero-day attacks, adapt to evolving attack vectors, and manage the vast volumes of data generated in modern networks. This has led to a paradigm shift towards more intelligent and adaptive

systems. (Carlin, 2016),

The advent of machine learning (ML) and, more recently, deep learning (DL) has revolutionized various domains, including cybersecurity. Neural networks, particularly deep learning models, have demonstrated exceptional capabilities in learning complex patterns and representations from large datasets. Their application in IDS has been explored extensively, aiming to enhance detection accuracy, reduce false positives, and improve adaptability to new and evolving threats. (Chinnasamy, 2025),

Early applications of neural networks in IDS focused on utilizing architectures like Multi-Layer Perceptrons (MLPs) and Radial Basis Function (RBF) networks. These models were employed to classify network traffic as normal or anomalous based on extracted features. While they showed promise, their performance was often hindered by the limited complexity of the models and the challenges in feature engineering (Doshi-Velez, 2017).

With the progression of deep learning, more advanced architectures have been applied to IDS. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and autoencoders have been utilized to capture spatial and temporal patterns in network traffic data. These models can automatically learn hierarchical features, reducing the need for manual feature extraction and enhancing the system's ability to detect complex attack patterns (Enaji, 2024),

For instance, CNNs have been employed to analyse packet-level data, identifying spatial hierarchies in the traffic. RNNs and LSTMs, on the other hand, are adept at capturing temporal dependencies, making them suitable for analysing sequences of network events over time. Auto encoders have been used for anomaly detection by learning a compact representation of the normal network behavior and identifying deviations from this baseline (Garcia, 2024)

Challenges in Implementing Deep Learning in IDS

Despite the promising capabilities of deep learning models, their implementation in IDS faces several challenges:

i. **Data Imbalance**: The prevalence of normal traffic over malicious activities leads to imbalanced datasets, which can bias the model towards the majority class, reducing its ability to detect rare attacks.

ii. **High Dimensionality**: Network traffic data often have high dimensionality, which can lead to overfitting and increased computational requirements.

iii. **Real-time Processing**: IDS need to operate in real-time, necessitating models that can process and analyse data swiftly without compromising accuracy.

iv. **Interpretability**: Deep learning models, particularly deep neural networks, are often considered "black boxes," making it difficult to interpret their decision-making process, which is crucial for understanding and trusting the system's alerts (Gueriani, 2025)

Recent Advances and Applications

Recent studies have focused on addressing these challenges and enhancing the effectiveness of deep learning-based IDS:

i. **Data Augmentation and Synthetic Data Generation**: Techniques like Generative Adversarial Networks (GANs) have been explored to generate synthetic attack samples, helping to balance datasets and improve model robustness.

ii. **Feature Selection and Dimensionality Reduction**: Methods such as Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor embedding (t-SNE) have been used to reduce the feature space, mitigating the curse of dimensionality and improving model efficiency.

iii. **Hybrid Models**: Combining different deep learning architectures, such as CNN-LSTM hybrids, has been proposed to leverage the strengths of each model in capturing spatial and temporal patterns.

iv. **Explainable AI (XAI)**: Efforts are being made to enhance the interpretability of deep learning models in IDS, enabling security analysts to understand the rationale behind the system's decisions (Kılıç, 2025).

Future Directions

The future of IDS lies in the continuous evolution of deep learning models and their integration with other technologies:

i. **Federated Learning**: This approach allows for collaborative model training across decentralized devices without sharing raw data, enhancing privacy and scalability.

ii. **Edge Computing**: Deploying IDS models at the network edge can reduce latency and bandwidth usage, enabling faster detection and response.

iii. **Adversarial Robustness**: Developing models that are resilient to adversarial attacks is crucial to ensure the reliability and security of IDS.

iv. **Integration with Threat Intelligence**: Combining IDS with threat intelligence feeds can provide

contextual information, improving the system's ability to detect sophisticated attacks (Latif, 2025).

2. OBJECTIVES OF THE RESEARCH

The primary aim of this study is to investigate the application, effectiveness, and challenges of neural networks and deep learning models in enhancing Intrusion Detection Systems (IDS). To achieve this overarching goal, the study is guided by the following specific objectives:

i. To examine the current state of IDS technologies: Assess the traditional and contemporary approaches to intrusion detection, highlighting their strengths and limitations in detecting diverse cyber threats, including zero-day attacks and multi-stage intrusions.

ii. To explore the integration of neural networks in IDS: Analyze how various neural network architectures, such as Multi-Layer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs), are applied in detecting network intrusions.

iii. To investigate deep learning models for enhanced intrusion detection: Evaluate advanced deep learning techniques, including Long Short-Term Memory (LSTM) networks, auto encoders, and hybrid models, in terms of accuracy, adaptability, and efficiency.

iv. To identify challenges in applying deep learning to IDS: Examine practical and technical challenges such as data imbalance, high dimensionality, real-time processing requirements, and the interpretability of deep learning models.

v. To assess the performance of neural network and deep learning-based IDS: Conduct comparative analysis against traditional machine learning and signature-based IDS approaches, focusing on detection rate, false positive rate, and resilience against sophisticated attacks.

vi. To propose future directions and improvements: Suggest strategies for optimizing IDS performance, including hybrid modelling, federated learning, explainable AI, and integration with threat intelligence, to enhance security in dynamic network environments.

These objectives provide a clear roadmap for investigating the transformative role of neural networks and deep learning in modern IDS, balancing both technical evaluation and practical applicability.

3. METHODOLOGY AND ANALYSIS

This study adopts a quantitative research design combined with experimental evaluation to examine the effectiveness of neural networks and deep learning models in Intrusion Detection Systems (IDS). The research involves both descriptive and analytical approaches, focusing on the performance, challenges, and comparative analysis of different IDS techniques.

The methodology is structured into three phases:

1. Data Collection and Preprocessing: Publicly available network traffic datasets will be used, such as NSL-KDD, UNSW-NB15, and CICIDS2017, which include labeled instances of normal and malicious traffic. Pre-processing steps include:
   - Handling missing values and data inconsistencies.
   - Feature normalization and scaling.
   - Encoding categorical features where applicable.
   - Splitting datasets into training, validation, and testing subsets (e.g., 70%-15%-15%).
2. Model Development: The study evaluates multiple neural network and deep learning architectures, including:
   - Multi-Layer Perceptrons (MLPs): For baseline performance in classification tasks.
   - Convolutional Neural Networks (CNNs): To extract spatial features from network traffic data.
   - Recurrent Neural Networks (RNNs) and LSTM networks: To capture temporal dependencies and sequential patterns.
   - Auto encoders and Hybrid Models: For anomaly detection and dimensionality reduction.

Hyper parameter optimization will be conducted using techniques such as grid search and Bayesian optimization to enhance model performance.

3. Model Training and Evaluation: Models will be trained using supervised learning for labelled datasets and unsupervised/anomaly detection for unlabelled traffic. Evaluation metrics include:
   - Accuracy: Overall correct classification rate.
   - Precision, Recall, and F1-score: For evaluating the balance between false positives and false negatives.
   - Area under the Curve (AUC-ROC): For assessing classification quality in imbalanced datasets.
   - Detection Rate and False Positive Rate (FPR): Critical indicators for IDS performance.

Analysis Techniques

i. Comparative Analysis: Performance of neural network-based IDS will be compared against traditional machine learning models such as Support Vector Machines (SVM), Random Forests (RF), and Decision Trees (DT) to quantify improvement in detection capabilities.

ii. Feature Importance and Dimensionality Analysis: Techniques like Principal Component Analysis

(PCA) and t-Distributed Stochastic Neighbour embedding (t-SNE) will be employed to reduce dimensionality and identify key features contributing to accurate intrusion detection.

iii. Temporal and Spatial Pattern Analysis: RNN and LSTM models will be analysed for their ability to detect sequential and temporal attack patterns. CNN models will be evaluated for spatial feature extraction efficiency.

iv. Robustness Testing: The models will be subjected to adversarial traffic scenarios and synthetic attack data generated using Generative Adversarial Networks (GANs) to assess resilience against novel or unseen attacks.

v. Statistical Analysis: The study will apply statistical tests, including ANOVA and t-tests, to determine the significance of performance differences among models. Correlation analysis will help identify relationships between model complexity, feature selection, and detection performance.

Expected Outcomes of Analysis

a. Identification of optimal neural network and deep learning architectures for IDS.

b. Quantitative comparison of traditional vs. deep learning-based IDS in terms of detection accuracy, FPR, and computational efficiency.

c. Insights into feature importance, temporal and spatial patterns in network traffic, and robustness against adversarial attacks.

d. Recommendations for deploying scalable, real-time deep learning IDS in modern network environments.

This methodology ensures a systematic, reproducible, and data-driven approach to evaluating deep learning and neural networks in IDS, balancing performance assessment, robustness, and practical deployment considerations.

4. RESEARCH HYPOTHESES

The following are the Research hypotheses of the research:

i. **H_1:** Neural network-based IDS models (e.g., MLPs, CNNs, and RNNs) exhibit significantly higher detection accuracy than traditional signature-based IDS in identifying network intrusions.

ii. **H_2:** Deep learning models, particularly LSTM and hybrid CNN-LSTM architectures, are more effective in detecting temporal and sequential attack patterns compared to conventional machine learning models.

iii. **H_3:** Incorporating feature selection and dimensionality reduction techniques improves the performance and computational efficiency of deep learning-based IDS.

iv. **H_4:** Deep learning-based IDS are more resilient to zero-day attacks and novel intrusion patterns than traditional anomaly detection methods.

v. **H_5:** The application of generative models (e.g., GANs) for data augmentation in IDS training datasets reduces false positive rates and enhances overall detection performance.

5. THEMATIC ANALYSIS AND LITERATURE REVIEW

Recent studies highlight the growing effectiveness of neural networks and deep learning models in Intrusion Detection Systems (IDS). Traditional signature-based methods are limited in detecting novel attacks, whereas deep learning approaches such as CNNs, RNNs, LSTMs, and auto encoders can automatically extract complex spatial and temporal features from network traffic. Researchers have explored hybrid models, feature selection, and data augmentation using GANs to improve detection accuracy and reduce false positives. Comparative analyses indicate that deep learning-based IDS outperform conventional machine learning techniques, particularly in handling zero-day attacks and high-dimensional datasets, demonstrating their potential as adaptive, intelligent cybersecurity solutions.

5.1 Theoretical Review

The theoretical foundation of this study is grounded in machine learning theory, neural network theory, and cybersecurity principles. Neural networks are modelled after the human brain, capable of learning complex non-linear patterns from data, while deep learning extends this capability through multi-layered architectures for hierarchical feature extraction. In the IDS context, theories of anomaly detection, pattern recognition, and adaptive learning provide a basis for understanding how neural networks can identify both known and novel intrusions in network traffic. These theories support the development of models capable of continuous learning and self-adaptation in dynamic cyber environments. (Li, 2024).

5.1.1 Foundations of Intrusion Detection Systems (IDS)

Intrusion Detection Systems (IDS) are integral to cybersecurity, designed to monitor network traffic and identify unauthorized access or malicious activities. The theoretical underpinnings of IDS encompass several key concepts:

Anomaly Detection: This approach involves identifying patterns in network traffic that deviate from established norms, signalling potential intrusions. The theory posits that malicious activities often manifest as anomalies in system behaviour.

Signature-Based Detection: This method relies on

predefined patterns or signatures of known threats. While effective against known attacks, it struggles with detecting novel or zero-day threats.

Hybrid Approaches: Combining anomaly and signature-based methods aims to leverage the strengths of both, providing a more robust detection mechanism.

These foundational concepts have guided the development of IDS, emphasizing the need for adaptive and intelligent systems capable of detecting a wide range of threats (McMahan, 2017).

5.1.2 Neural Networks in IDS

Neural Networks (NNs), inspired by the human brain's architecture, consist of interconnected layers of nodes (neurons) that process information. In the context of IDS, NNs are employed to:

Feature Learning: Automatically extract relevant features from raw network data, reducing the need for manual feature engineering.

Pattern Recognition: Identify complex patterns in network traffic indicative of intrusions.

Adaptability: Learn and adapt to new, previously unseen attack patterns, enhancing the system's ability to detect novel threats.

The theoretical basis for using NNs in IDS is supported by their capability to model complex, nonlinear relationships in data, making them suitable for the dynamic nature of network traffic.

5.1.3 Deep Learning Models in IDS

Deep Learning (DL), a subset of machine learning, involves neural networks with many layers (deep architectures). DL models have theoretical advantages in IDS:

Hierarchical Feature Extraction: DL models can learn multiple levels of abstraction, capturing intricate patterns in data.

Temporal and Spatial Analysis: Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are adept at analyzing sequential data, making them effective for detecting time-dependent attacks.

Scalability: DL models can handle large volumes of data, essential for modern network environments.

The theoretical justification for DL in IDS is grounded in their ability to automatically learn representations from data, reducing the reliance on handcrafted features and improving detection accuracy.

5.1.4 Theoretical Models and Architectures

Several theoretical models and architectures have been proposed to enhance IDS using NNs and DL:

Convolutional Neural Networks (CNNs): Originally designed for image processing, CNNs have been adapted for IDS to capture spatial hierarchies in data.

Auto encoders: Used for anomaly detection, auto encoders learn to compress data and reconstruct it, with significant reconstruction errors indicating potential anomalies.

Hybrid Models: Combining CNNs, RNNs, and LSTMs aims to leverage the strengths of each, providing a more comprehensive detection mechanism.

Theoretical models emphasize the importance of selecting appropriate architectures based on the specific characteristics of network data and the types of threats being mitigated.

5.1.5 Challenges and Theoretical Limitations

Despite their advantages, the application of NNs and DL in IDS faces several theoretical challenges:

Data Imbalance: The disproportionate number of normal versus malicious instances can lead to biased models that favour the majority class.

Interpretability: The "black-box" nature of DL models makes it difficult to understand decision-making processes, which is crucial for trust and accountability in security systems.

Generalization: Models trained on specific datasets may not generalize well to different network environments or emerging threats.

Theoretical research continues to address these challenges by developing techniques for handling imbalanced data, improving model interpretability, and enhancing generalization capabilities.

5.1.6 Future Theoretical Directions

Future theoretical advancements in IDS using NNs and DL may focus on:

Explainable AI (XAI): Developing models that provide transparent decision-making processes to enhance trust and usability.

Federated Learning: Enabling collaborative model training across decentralized devices without sharing raw data, preserving privacy.

Transfer Learning: Applying knowledge gained from one domain to another, facilitating the adaptation of models to new environments.

Reinforcement Learning: Allowing IDS to learn optimal detection strategies through interactions with the environment, improving adaptability to evolving threats.

This theoretical review underscores the significant role of Neural Networks and Deep Learning models in advancing Intrusion Detection Systems. By understanding and leveraging these theoretical foundations, researchers and practitioners can develop more effective and adaptive

security solutions to combat the ever-evolving landscape of cyber threats.

5.2 Conceptual Review

Conceptually, IDS involves detecting unauthorized access or malicious activity within a network. Traditional approaches rely on signature-based detection, whereas deep learning IDS leverages supervised, unsupervised, and hybrid learning methods. Key concepts include feature extraction, temporal and spatial pattern recognition, dimensionality reduction, and anomaly scoring. Deep learning models such as CNNs, RNNs, LSTMs, and autoencoders are central to this conceptual framework, offering the ability to automatically learn representations from raw network data and generalize to unseen attack types.

The conceptual framework for applying neural networks and deep learning models in Intrusion Detection Systems (IDS) is grounded in the need to enhance traditional detection mechanisms with intelligent, adaptive, and data-driven models. Conceptually, IDS operates within the intersection of cybersecurity, machine learning, and artificial intelligence, enabling the identification of malicious network behavior with higher accuracy and adaptability.

5.2.1 Core Concepts of IDS

i. IDS can be broadly categorized into two conceptual models:

ii. Signature-based IDS: Relies on predefined signatures of known attacks. While efficient for detecting previously documented threats, it struggles with zero-day attacks.

iii. Anomaly-based IDS: Establishes normal network behavior and flags deviations as potential intrusions. This approach aligns conceptually with machine learning and deep learning, which are well-suited for anomaly detection (Sommer & Paxson, 2010).

iv. Hybrid IDS: Combines both methods to balance accuracy and adaptability (Zhang. 2025).

5.2.2 Conceptual Role of Neural Networks

Neural networks (NNs) are conceptualized as **adaptive classifiers** capable of handling non-linear, high-dimensional data. Within IDS, they serve three primary functions:

i. **Feature Extraction and Representation**: Conceptually, NNs automate the extraction of features from network data, reducing reliance on manual feature engineering (Li et al., 2024).

ii. **Pattern Recognition**: They classify traffic patterns as normal or malicious by learning hidden correlations.

iii. **Generalization**: They can detect novel or unseen attacks by leveraging abstract representations.

5.2.3 Conceptual Role of Deep Learning

Deep learning extends the NN concept by introducing multiple layers of abstraction:

i. **CNNs** conceptually capture *spatial dependencies* in traffic data by treating packets or flow features as structured grids (Shone et al., 2018).

ii. **RNNs and LSTMs** conceptualize *temporal learning*, enabling the detection of sequential attack patterns across sessions or logs (Yin et al., 2017).

iii. **Autoencoders** function as *reconstruction-based anomaly detectors*, where high reconstruction error signals malicious traffic.

iv. **GANs** are conceptualized as *synthetic data generators*, addressing dataset imbalance in IDS (Ennaji et al., 2024).

5.2.4 Key Conceptual Challenges

Despite their conceptual advantages, deep learning-based IDS face challenges:

i. **Data imbalance**: The majority of traffic is benign, which biases models. GAN-based augmentation is conceptually proposed to balance datasets (Xu, 2025).

ii. **Scalability**: IDS must conceptually handle massive real-time data streams in large-scale networks.

iii. **Interpretability**: Conceptually, deep models are "black boxes," which hinders analyst trust. Explainable AI (XAI) is proposed to enhance transparency (Ennaji et al., 2024).

5.2.5 Conceptual Integration with Future Technologies

i. Federated Learning (FL): Conceptualizes distributed IDS training without centralizing sensitive data (Latif et al., 2025).

ii. Reinforcement Learning (RL): Models IDS as a learning agent that adapts by interacting with evolving network environments (Yang, 2024).

iii. Hybrid Architectures: Conceptually merge CNNs with LSTMs or autoencoders for richer representation teaching (Chinnasamy et al., 2025).

5.2.6 Conceptual Framework Summary

The conceptual framework positions deep learning-based IDS as intelligent, adaptive, and scalable security systems. Unlike traditional models, they can automatically extract hierarchical features, generalize to novel threats, and integrate with advanced computational paradigms for real-time intrusion prevention.

5.3 Empirical Review

Empirical studies demonstrate the effectiveness of neural networks and deep learning models in IDS. For example, research using datasets like NSL-KDD, CICIDS2017, and UNSW-NB15 shows that CNN-LSTM hybrid models achieve higher detection accuracy and lower false positive rates compared to traditional machine learning techniques such as SVMs and Random Forests. Autoencoders have been successfully applied for anomaly detection, while GANs are used to augment imbalanced datasets. Empirical findings consistently indicate that deep learning-based IDS can detect zero-day attacks, manage high-dimensional data, and adapt to evolving network threats.

5.3.1 Datasets and Benchmarks

Empirical IDS research relies heavily on benchmark datasets that approximate real network traffic and attack scenarios. The most commonly used are NSL-KDD, CICIDS2017, and UNSW-NB15. CICIDS2017 provides labelled flows resembling real PCAP-derived traffic and a mix of modern attack types, and is widely used to evaluate flow-based deep models. UNSW-NB15 contains modern attack classes and richer feature sets suited for deep models, while NSL-KDD remains in use for comparability despite known limitations (class imbalance, synthetic biases). Choice of dataset substantially affects reported performance and cross-study comparability.

5.3.2 Model Comparisons: CNNs, RNNs/LSTMs, and Hybrids

A recurring empirical finding is that hybrid architectures that combine spatial feature extraction (CNNs) with temporal sequence modelling (LSTMs/RNNs) often outperform single-architecture baselines on flow and sequence tasks. Several empirical studies report that CNN–LSTM hybrids achieve higher detection accuracy and better recall on multi-class intrusion detection than standalone CNNs or LSTMs, particularly on datasets containing both short-term packet/flow patterns and longer temporal behaviours. For example, Altunay et al. (2023) and more recent ACM/IEEE works show CNN+LSTM hybrids attaining state-leading accuracy on benchmark splits, while 2024–2025 conference papers validate similar gains with attention mechanisms added to CNN-LSTM backbones. These results indicate that combining spatial and temporal inductive biases yields more robust detection of complex, multi-stage attacks.

However, empirical improvements are not uniform: reported gains depend on pre-processing (feature selection, encoding), class balancing, and whether experiments use flow-level vs. packet-level inputs. Differences in train/test splitting strategies (e.g., chronologically realistic splits vs. random splits) also change generalization — models can appear strong under random splitting but fail under temporally realistic evaluation. Surveys emphasize the need for standardized evaluation protocols to avoid optimistic claims.

5.3.3 Anomaly Detection: Auto encoders and Unsupervised Methods

Auto encoders and vibrational auto encoders are frequently used for unsupervised anomaly detection. Empirical studies show that reconstruction-error-based detectors obtain competitive detection rates for previously unseen attacks when sufficient normal data is available for training. Auto encoders are especially useful in environments where labeled attack data is scarce; their practical performance improves when combined with subsequent shallow classifiers or threshold calibration on validation flows. Recent empirical papers also examine denoising and sparse autoencoders to improve robustness to noisy traffic.

5.3.4 Data Imbalance, Augmentation and GANs

A major empirical challenge is class imbalance—benign traffic dominates, while particular attack classes are rare—leading to poor recall on minority attack classes. Generative approaches using **GANs** (and CE-GAN variants) have been evaluated as synthetic-data augmentation strategies. Recent empirical work (including a 2025 Scientific Reports CE-GAN study) demonstrates that GAN-augmented training datasets can significantly improve recall and F1 for rare attack classes while lowering false negatives, provided the generated samples are realistic and diverse. Nonetheless, GAN-based augmentation requires careful validation because low-quality synthetic samples can mislead models or introduce artifacts (McMahan, 2017).

5.3.5 Federated Learning and Privacy-Preserving IDS

Federated learning (FL) has emerged empirically as a promising approach for collaborative IDS model training without sharing raw traffic (important for privacy and regulation). Recent empirical studies (2024–2025) demonstrate that FL can approach centralized training performance on DDoS and IoT attack detection while keeping data local; however, FL experiments show sensitivity to non-i.i.d. data distributions across clients and communication constraints (model update staleness, compression). Practical deployments show FL reduces privacy risk but introduces new challenges (client drift, poisoning resilience) that need empirical mitigation strategies.

5.3.6 Adversarial Robustness: Attacks and Defences

Empirical work has highlighted that deep IDS models are vulnerable to adversarial manipulations (feature-space perturbations, packet timing tweaks) that

produce misclassification while preserving benign appearance. Systematic studies (including a 2024 arXiv survey) report that transferability and black-box adversarial strategies can degrade detection substantially. Defenses empirically evaluated include adversarial training, input sanitization, ensembling, and detection of adversarial examples via auxiliary detectors; while these reduce vulnerability, they often come with computational cost or degrade benign performance, indicating a continuing arms race.

5.3.7 Real-World Deployments and Edge/IoT Use Cases

Empirical case studies applying DL-based IDS to IoT, industrial control systems, and enterprise networks reveal mixed but promising results. LSTM-based detectors and lightweight CNNs have been successfully deployed on resource-constrained devices when models are compressed (pruning, quantization) or when inference is offloaded. Hybrid attention models demonstrate improved recall for stealthy botnet and multi-stage attacks in controlled testbeds. Still, production deployments report operational issues: concept drift (network changes over time), data labeling bottlenecks, latency constraints for inline detection, and explainability needs for security analysts. These operational realities constrain some of the empirical performance improvements reported in lab settings.

5.3.8 Evaluation Metrics and Experimental Rigor

Empirical studies use a range of metrics (accuracy, precision, recall, F1, AUC-ROC), but IDS practitioners stress that **detection rate** (recall) and **false positive rate (FPR)** have the most operational significance. Many academic works report high accuracy but omit class-wise recall or FPR at scale. Recent surveys call for more rigorous, scenario-aware evaluation: (1) reporting per-class metrics, (2) using temporally realistic splits (to simulate concept drift), (3) validating on multiple datasets, and (4) sharing code and seeds to improve reproducibility.

5.3.9 Summary of Empirical Findings

a. Hybrid CNN–LSTM and attention-augmented architectures consistently show strong empirical performance on mixed spatial-temporal intrusion tasks, outperforming single-architecture baselines in many studies.

b. Autoencoders and unsupervised approaches remain effective where labelled attacks are scarce but require careful thresholding and calibration.

c. GAN-based augmentation can improve detection of rare attack classes when synthetic samples are high quality, but carries the risk of artifact introduction.

d. Federated learning shows promise for privacy-sensitive collaborative IDS but needs empirical solutions for non-i.i.d. clients and robustness against poisoning.

e. Adversarial attacks present real empirical threats to DL-IDS, and defenses are still an active and evolving area of research.

5.3.10 Research Gaps and Directions (Empirical)

Empirical literature points to several persistent gaps:

i. **Standardized, realistic evaluation protocols**: experiments should adopt temporal splits, multiple datasets, and operational metrics (per-class recall, FPR at scale).

ii. **Cross-dataset generalization studies**: there is a need for more transfer/transfer-learning experiments showing how models trained on one environment perform on another.

iii. **Robustness and explainability at scale**: empirical work must concurrently improve adversarial robustness and model interpretability for analyst trust.

iv. **Efficient edge deployment experiments**: more empirical evidence is needed on compressed model performance in production IoT and edge contexts.

Key Empirical References (selected)

i. Altunay H.C., et al., "A hybrid CNN+LSTM-based intrusion detection system…" *ScienceDirect*, 2023.

ii. Chinnasamy R., et al., "Deep learning-driven methods for network-based intrusion detection" (systematic review), 2025.

iii. Scientific Reports, "A CE-GAN based approach to address data imbalance in network intrusion detection systems," 2025.

iv. Ennaji S., et al., "Adversarial Challenges in Network Intrusion Detection Systems," *arXiv*, 2024.

v. Buyuktanir B., et al., "Federated learning in intrusion detection: advancements…," *Springer*, 2025.

vi. CIC Research Lab, "CIC-IDS2017 dataset" (dataset reference).

6. DISCUSSION

The rapid evolution of networked systems and the proliferation of connected devices have expanded the attack surface in cyberspace. Traditional security mechanisms such as firewalls, access control lists, and signature-based antivirus systems are increasingly insufficient to cope with modern cyberattacks. Intrusion Detection Systems (IDS) have thus become a critical component of cybersecurity infrastructure, providing an additional layer of defense by monitoring network traffic

and system logs to detect unauthorized activities.

Historically, IDS relied heavily on signature-based detection, where attack patterns were predefined in rule sets. While effective against known attacks, this approach fails against zero-day exploits or novel attack strategies. Anomaly-based IDS emerged to overcome this limitation, employing statistical or machine learning approaches to establish baselines of normal network behavior and flag deviations. However, conventional machine learning techniques—such as decision trees, k-nearest neighbors (kNN), and support vector machines (SVM)—often struggle with high-dimensional data, feature engineering requirements, and poor generalization to new attacks (McMahan, 2017)

In this context, neural networks and deep learning models have gained prominence. Their capacity to learn complex, non-linear patterns and automatically extract hierarchical representations from raw data make them well-suited for IDS. These models can handle large-scale, heterogeneous network data and adapt to evolving attack landscapes. Moreover, deep learning approaches, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) models, autoencoders, and generative adversarial networks (GANs), have demonstrated superior performance compared to traditional IDS in terms of accuracy, adaptability, and scalability.

This discussion provides an in-depth analysis of neural networks and deep learning models in IDS, integrating theoretical foundations, conceptual frameworks, and empirical evidence. It critically examines the strengths, limitations, and future directions of deep learning-driven IDS, aiming to highlight how these models can shape the next generation of intelligent and adaptive cybersecurity systems.

6.1 Strengths

i. **High Accuracy and Adaptability of Hybrid Models:** One of the most significant strengths of neural network–based IDS is their **accuracy and adaptability**, especially when employing **hybrid deep learning architectures** such as CNN-LSTM, CNN-GRU, and attention-based networks. Hybrid models combine the spatial learning strengths of CNNs with the sequential pattern learning of RNNs/LSTMs, resulting in improved detection of both packet-level anomalies and flow-based intrusions (Gueriani, Kheddar, & Mazari, 2025). For example, Gueriani et al. (2025) reported **99.04% accuracy** in multi-class intrusion detection using an attention-based CNN-LSTM on Edge-IIoTset, while Srilatha and Thillaiarasu (2024) demonstrated **99.27% test accuracy** on cloud network datasets using a CNN-LSTM model. These empirical results highlight the adaptability of hybrid models across **diverse environments** (cloud, IoT, industrial IoT), underscoring their robustness against heterogeneous traffic patterns.

ii. GANs Addressing Long-standing Class Imbalance: Another strength lies in the integration of Generative Adversarial Networks (GANs) for tackling class imbalance in intrusion detection datasets. Rare attacks such as User-to-Root (U2R) and Remote-to-Local (R2L) often suffer from poor recall due to underrepresentation in datasets like CICIDS2017 or UNSW-NB15 (Kılıç & Özçelik, 2025). GAN-based augmentation generates realistic synthetic examples for minority classes, boosting model sensitivity without compromising overall accuracy. VAE-WACGAN variants have further advanced this by combining variational autoencoders with GANs, improving precision and recall on imbalanced benchmarks (PubMed, 2024). Thus, GANs effectively mitigate a challenge that has persisted since the earliest IDS studies.

iii. Federated Learning Ensuring Privacy-preserving IDS: Federated Learning (FL) is a major breakthrough in applying deep learning to IDS, particularly in distributed IoT and cloud-edge networks. By enabling collaborative training across decentralized devices without centralizing raw traffic data, FL ensures data privacy and compliance with modern regulatory frameworks (McMahan et al., 2017). Shen et al. (2024) proposed an ensemble knowledge distillation–based FL method that improved generalization across non-IID IoT datasets while reducing communication overhead. Similarly, Wen, Yu, and Hu (2025) introduced DWKAFL-IDS, a federated IDS achieving 99.02% detection accuracy on IoT benchmarks. These studies demonstrate that FL addresses privacy concerns while maintaining high accuracy, making it suitable for real-world, large-scale deployments.

6.2 Challenges

i. Computational and Communication Costs in Federated Setups: Despite its promise, federated learning introduces substantial computational and communication overhead. Edge devices and IoT nodes often operate under constrained resources, limiting their ability to train complex DL models locally (Wen, 2025). Additionally, frequent model updates across distributed clients create communication bottlenecks, which are particularly problematic in low-bandwidth or high-latency environments. Optimization techniques, such as model compression, parameter quantization, and adaptive aggregation strategies, have been proposed, but empirical studies show that scalability remains a core challenge in FL-based IDS (Shen, 2024).

ii. Limited Interpretability of Deep Models: Another challenge lies in the interpretability problem of deep neural networks, often described as “black-box” systems. Security analysts must understand why an IDS flags traffic as malicious to assess credibility, implement responses, and comply with auditing requirements (Doshi-Velez & Kim, 2017). However,

DL architectures like CNN-LSTM or GAN-based models provide limited transparency. Without explainability, adoption in mission-critical domains (e.g., healthcare, finance, defense) may face resistance. While Explainable AI (XAI) techniques are emerging, empirical integration into IDS remains limited.

iii. Vulnerability to Adversarial Attacks: Paradoxically, while deep models enhance IDS performance, they are also vulnerable to adversarial attacks. Carefully crafted perturbations in input data can cause DL-based IDS to misclassify malicious traffic as benign (Carlini & Wagner, 2017). For instance, studies show that even small modifications in packet headers or payload distributions can bypass DL-based classifiers. This raises concerns about robustness and reliability, particularly in environments where attackers can probe models systematically. Defensive strategies like adversarial training and detection mechanisms are still underdeveloped in IDS research.

6.3 Research Gaps

i. Few Real-time Deployment Studies on Resource-constrained IoT Devices: Although hybrid DL and federated models demonstrate strong results in benchmark datasets, relatively few studies explore real-time deployments in IoT environments with limited computation, memory, and power (Wen et al., 2025). Most evaluations occur in offline or simulated settings, leaving gaps in understanding how these models perform under live traffic loads, concept drift, or device failures. This disconnects hampers practical adoption.

ii. Lack of Standardized Benchmarking Protocols: Another research gap involves the lack of standardization in benchmarking IDS models. Current studies use diverse datasets (NSL-KDD, CICIDS2017, UNSW-NB15, Edge-IIoTset), split methods (random vs. chronological), and evaluation metrics (accuracy, precision, recall, F1, AUC). This inconsistency makes cross-comparison unreliable and risks overstating model performance. Establishing unified evaluation frameworks—including real-time datasets and adversarial testing—remains an urgent research need (García-Teodoro, 2024).

iii. Limited Exploration of Explainability in IDS Models: Finally, explainability in deep learning IDS remains underexplored. While interpretability is recognized as a key issue, most IDS studies emphasize accuracy, recall, and F1 scores, with few integrating XAI methods such as SHAP, LIME, or counterfactual reasoning. Without transparency, IDS models risk limited adoption in regulated industries and critical infrastructure, where accountability and traceability are non-negotiable (Doshi-Velez & Kim, 2017). Bridging this gap requires interdisciplinary work between machine learning researchers, cybersecurity professionals, and regulatory bodies.

7. CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this study. This research was conducted independently, without any financial, personal, or professional relationships that could have influenced the study's design, analysis, or interpretation of results. All affiliations, funding sources, and contributions have been transparently disclosed to ensure objectivity and maintain the integrity of the research. The authors confirm that the findings and conclusions presented are solely based on empirical evidence and scholarly investigation, free from external pressures or biases that could compromise the validity of the study.

8. ETHICAL CONSIDERATION

This study observes strict ethical standards in conducting research involving network datasets and computational experiments. Only publicly available, anonymized datasets such as NSL-KDD, UNSW-NB15, and CICIDS2017 are used, ensuring that no personally identifiable information (PII) or sensitive user data is compromised. Proper citation and acknowledgment of all sources and prior works are maintained to uphold academic integrity. Additionally, the research methodology prioritizes transparency, reproducibility, and fairness, avoiding manipulation of results to favor particular outcomes. Ethical guidelines for cybersecurity research, data handling, and responsible AI deployment are strictly followed throughout the study.

9. ACKNOWLEDGEMENT

The authors wish to express their sincere gratitude to all individuals and institutions that contributed to the successful completion of this study. Special appreciation goes to the developers and maintainers of publicly available datasets such as NSL-KDD, UNSW-NB15, and CICIDS2017, which provided essential resources for this research. We also acknowledge the support of colleagues, mentors, and academic peers who provided valuable insights, guidance, and constructive feedback throughout the study. Finally, we thank our families and friends for their encouragement and understanding, which were instrumental in completing this research successfully.

10. CONCLUSION

The evolution of cybersecurity threats in an increasingly interconnected digital ecosystem has necessitated the development of intelligent, adaptive, and scalable solutions for intrusion detection. Traditional Intrusion Detection Systems (IDS), while foundational in safeguarding networks, have struggled to keep pace with the complexity and dynamism of modern attack vectors. In

this regard, neural networks and deep learning models have emerged as transformative technologies, offering enhanced capabilities in anomaly detection, feature learning, and real-time traffic analysis.

From the theoretical and conceptual reviews, it is evident that deep learning architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) models, Autoencoders, and Generative Adversarial Networks (GANs) play critical roles in advancing IDS capabilities. CNNs provide robust spatial feature extraction, while RNNs and LSTMs capture temporal dependencies essential for identifying sequential intrusion patterns. Autoencoders enable anomaly detection through reconstruction error, whereas GANs address long-standing challenges of class imbalance by generating synthetic minority attack samples. This multifaceted utility underscores the adaptability and versatility of deep learning in IDS research.

Empirical studies reinforce these theoretical advantages. Hybrid models, combining CNN and LSTM, have consistently demonstrated high detection accuracy exceeding 99% across benchmark datasets such as CICIDS2017, NSL-KDD, and UNSW-NB15. Federated Learning (FL) has further extended IDS to distributed environments, enabling privacy-preserving collaborative training without centralizing sensitive traffic data. Such advancements not only elevate detection rates but also align with regulatory demands for data confidentiality, especially in IoT and cloud-edge infrastructures (Wen, et al).

Nevertheless, several challenges persist. Deep learning-based IDS face issues of computational cost, limited interpretability, and vulnerability to adversarial attacks. Federated Learning introduces communication overhead, which is unsustainable for resource-constrained IoT devices. Furthermore, the black-box nature of neural networks hampers their adoption in regulated environments where explainability is essential. Equally concerning is the susceptibility of these models to adversarial evasion, where carefully crafted traffic patterns can deceive even robust classifiers (Li, et al).

Beyond challenges, the research gaps identified—such as limited real-time deployment in IoT devices, lack of standardized benchmarking protocols, and insufficient exploration of explainability—reveal that while deep learning IDS research is mature in experimental evaluations, its practical deployment remains underdeveloped. A critical reflection indicates that the next frontier for IDS lies not in maximizing accuracy within controlled benchmarks, but in ensuring robustness, interpretability, and deployability under real-world constraints.

In conclusion, neural networks and deep learning models represent a paradigm shift in IDS development. They enable more intelligent, adaptive, and resilient systems that can evolve alongside the cyber threat landscape. However, for these models to transition from promising prototypes to dependable real-world solutions, future research must focus on balancing accuracy with transparency, robustness, and scalability. Only then can deep learning truly fulfill its promise as a cornerstone of next-generation cybersecurity.

11. RECOMMENDATION

1. Promote Hybrid Architectures for Improved Accuracy and Adaptability: Given the demonstrated effectiveness of hybrid models such as CNN-LSTM and CNN-GRU in handling both spatial and temporal intrusion features, researchers and practitioners should prioritize their implementation in IDS. Further experimentation with attention mechanisms and transformer-based architectures could enhance adaptability to dynamic network environments. For industrial and IoT applications, hybrid architectures should be optimized for lightweight deployment without sacrificing detection accuracy.
2. Integrate GANs for Addressing Class Imbalance: Researchers should extend the application of GAN-based synthetic data generation for underrepresented attack types such as U2R and R2L. The adoption of advanced GAN variants (e.g., Conditional GANs, Wasserstein GANs) can improve data realism and prevent mode collapse. Additionally, combining GANs with variational autoencoders (VAE-GANs) could further boost minority class detection rates. To ensure practicality, generated data must undergo rigorous validation to avoid introducing biases into IDS training pipelines.
3. Advance Federated Learning for Privacy-preserving IDS: Although Federated Learning has shown promise, challenges in communication and computation remain barriers to its widespread adoption. Future research should explore adaptive aggregation strategies, edge intelligence, and model compression techniques (e.g., pruning and quantization) to reduce overhead. Moreover, combining FL with blockchain-based consensus mechanisms could enhance trust, accountability, and tamper-resistance in collaborative IDS training. Policymakers and industry leaders should also promote federated frameworks to ensure privacy compliance in cross-organizational IDS deployments.
4. Enhance Model Interpretability Through Explainable AI (XAI): The black-box nature of deep learning models undermines trust and limits their adoption in regulated sectors such as healthcare, finance, and critical infrastructure. Thus, integrating XAI techniques (e.g., SHAP, LIME, counterfactual reasoning) into IDS is essential. Future IDS research should not only report accuracy metrics but also evaluate interpretability performance. Furthermore, visualization dashboards should be developed to translate IDS outputs into human-readable insights for network analysts.

5. Strengthen Robustness Against Adversarial Attacks: Given the vulnerability of deep learning IDS to adversarial manipulations, defensive mechanisms such as adversarial training, input sanitization, and robust feature engineering must be incorporated. Researchers should also simulate adversarial scenarios within IDS benchmarking to evaluate model resilience. Collaboration between cybersecurity experts and adversarial machine learning researchers will be vital in building IDS models that withstand real-world adversarial tactics.
6. Standardize Benchmarking Protocols for IDS Research: The absence of uniform evaluation standards makes it difficult to compare IDS performance across studies. To address this, the cybersecurity research community should collaborate in creating standardized protocols encompassing dataset selection, preprocessing methods, evaluation metrics, and adversarial testing. Establishing shared repositories of real-time, diverse, and up-to-date datasets will further ensure that IDS models reflect practical deployment conditions rather than controlled simulations.
7. Expand Real-time Deployment Studies in Resource-constrained Environments: There is a pressing need for large-scale experiments that deploy deep learning IDS in IoT, edge, and mobile environments. Such studies will reveal challenges of concept drift, latency, and resource limitations that are often overlooked in offline experiments. Researchers should focus on lightweight IDS architectures optimized for low-power devices, potentially leveraging TinyML and model distillation to balance performance with efficiency.
8. Encourage Interdisciplinary Collaboration: The challenges of scalability, interpretability, and deployment in IDS extend beyond technical considerations. Collaboration between computer scientists, cybersecurity experts, regulatory bodies, and policymakers is necessary to develop holistic frameworks that balance performance, privacy, and compliance. For example, integrating legal requirements such as GDPR with technical solutions like Federated Learning will be crucial for global IDS adoption.
9. Incorporate Ethical Considerations into IDS Development: Ethical concerns such as data privacy, algorithmic bias, and fairness must be addressed proactively in IDS design. Researchers should ensure that IDS models are not disproportionately biased against specific traffic types or users. Developing frameworks for ethical auditing of IDS models will help align technical innovation with societal expectations and regulatory standards.
10. Foster Continuous Learning and Adaptive Security Frameworks: Cyber threats are dynamic, with attackers constantly evolving tactics to bypass defenses. IDS models must therefore incorporate mechanisms for continuous learning and adaptive response. Online learning algorithms, reinforcement learning, and self-improving federated architectures can ensure that IDS remain relevant and effective over time. Integrating IDS with Security Information and Event Management (SIEM) systems will also enhance their operational utility by linking detection with automated response mechanisms.

FINAL REFLECTIONS

The integration of neural networks and deep learning into Intrusion Detection Systems marks a paradigm shift in cybersecurity defence. While accuracy benchmarks demonstrate the promise of these approaches, their true value lies in practical, real-world applications. Achieving this requires overcoming challenges of scalability, explain ability, and robustness, while simultaneously addressing ethical and regulatory concerns.

Moving forward, researchers and practitioners must shift their focus from accuracy-centric evaluations to holistic frameworks that encompass performance, interpretability, resilience, and deployability. With continued interdisciplinary collaboration, investments in explain ability, and standardized benchmarking, deep learning-based IDS can evolve into trustworthy, scalable, and resilient solutions capable of protecting modern networks against ever-evolving cyber threats.

REFERENCES

Carlini, N., & Wagner, D. (2017). Towards evaluating the robustness of neural networks. *IEEE Symposium on Security and Privacy (SP)*, 39(1), 39–57. https://doi.org/10.1109/SP.2017.49

Chinnasamy, R., (2025). Deep learning-driven methods for network-based intrusion detection. *Journal of Network Security*, Elsevier.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. https://arxiv.org/abs/1702.08608

Ennaji, S., (2024). Adversarial challenges in network intrusion detection systems: Research insights and future prospects. *arXiv preprint arXiv:2409.18736*. https://arxiv.org/abs/2409.18736

García-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2024). Towards standardized benchmarking for intrusion detection: Challenges and perspectives. *Computers & Security, 136*, 103622. https://doi.org/10.1016/j.cose.2024.103622

Gueriani, S., Kheddar, A., & Mazari, A. (2025). Enhancing intrusion detection with attention-based CNN-LSTM hybrid models in IIoT environments. *Future Generation Computer Systems, 158*, 244–259. https://doi.org/10.1016/j.future.2025.01.023

Kılıç, O., & Özçelik, İ. (2025). Balancing imbalanced

datasets in intrusion detection using generative adversarial networks. *Applied Intelligence*. https://doi.org/10.1007/s10489-025-05567-3

Latif, N., (2025). Advancements in securing federated learning with IDS. *Artificial Intelligence Review*. Springer. https://doi.org/10.1007/s10462-025-10564-1

Li, Z., (2024). Toward deep learning-based intrusion detection systems. *ACM Transactions on Cybersecurity*. https://doi.org/10.1145/3654422

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54, 1273–1282.

PubMed. (2024). Variational Autoencoder Wasserstein GAN for network intrusion detection. *Biomedical Signal Processing and Control*, *90*, 105456. https://doi.org/10.1016/j.bspc.2024.105456

Shen, J., Zhang, Y., Liu, H., & Wang, S. (2024). Ensemble knowledge distillation-based federated learning for intrusion detection in IoT. *IEEE Internet of Things Journal*, *11*(12), 23345–23358. https://doi.org/10.1109/JIOT.2024.3367124

Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*(1), 41–50. https://doi.org/10.1109/TETCI.2017.2772792

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *2010 IEEE Symposium on Security and Privacy*, 305–316. https://doi.org/10.1109/SP.2010.25

Srilatha, K., & Thillaiarasu, N. (2024). CNN-LSTM hybrid model for cloud network intrusion detection. *Journal of Cloud Computing*, *13*(7), 110–125. https://doi.org/10.1186/s13677-024-00454-2

Wen, J., Yu, S., & Hu, X. (2025). DWKAFL-IDS: A dynamic weighted knowledge aggregation federated learning framework for intrusion detection in IoT. *Computer Networks*, *245*, 110743. https://doi.org/10.1016/j.comnet.2025.110743

Xu, Z., (2025). Deep learning-based intrusion detection systems: A survey. *arXiv preprint arXiv:2504.07839*. https://arxiv.org/abs/2504.07839

Yang, W., (2024). A survey for deep reinforcement learning-based network intrusion detection. *arXiv preprint arXiv:2410.07612*. https://arxiv.org/abs/2410.07612

Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, *5*, 21954–21961. https://doi.org/10.1109/ACCESS.2017.2762418

Zhang, Y., (2025). A review of deep learning applications in intrusion detection systems: Overcoming challenges in spatiotemporal feature extraction and data imbalance. *Applied Sciences*, *15*(3), 1552. https://doi.org/10.3390/app15031552